

Quadratic Polynomials

Polynomials in One Variable.

After studying linear functions $y = ax + b$, the next step is to study *quadratic polynomials*, $y = ax^2 + bx + c$, whose graphs are parabolas. Initially one studies the simpler special case

$$y = ax^2 + c \quad (1)$$

If $a > 0$ these parabolas have a minimum at $x = 0$ and open upward, while if $a < 0$ they have a maximum at $x = 0$ and open downward.

One can reduce the more general quadratic polynomial

$$y = ax^2 + bx + c \quad (2)$$

to the special case (1) by a change of variable, $x = v + r$ translating x by r , where r is to be found. Substituting this into (2) we find

$$y = a(v + r)^2 + b(v + r) + c = av^2 + (b + 2ar)v + ar^2 + br + c.$$

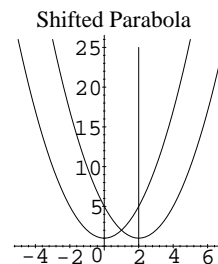
We now pick r to remove the linear term in v , that is, $b + 2ar = 0$ so $r = -b/(2a)$. Then

$$y = av^2 + k = a(x - r)^2 + k, \quad (3)$$

where $k = c - b^2/(4a)$. Thus, $x = r$ is the *axis of symmetry* of this parabola.

This procedure is equivalent to “completing the square”, a procedure that should be familiar from algebra. Another way to find r is to observe that the only critical point of (2) is at $x = -b/(2a)$. Thus translating by $b/(2a)$ places this critical point on the vertical axis.

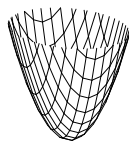
EXAMPLE. On the right are the graphs of $y = x^2 + 1$ and $y = (x - 2)^2 + 1 = x^2 - 2x + 2$. They clearly show the graph on the right is merely a translation of the graph on the left.



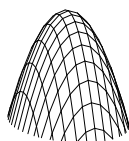
Polynomials in Several Variables.

MAXIMA, MINIMA, AND SADDLE POINTS.

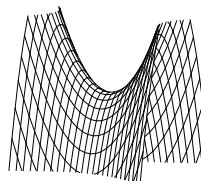
There are more interesting possibilities for quadratic polynomials in two variables.



$$w = 2x^2 + y^2$$



$$w = -(2x^2 + y^2)$$



$$w = -2x^2 + y^2$$

From the graphs it is clear that the first has a *minimum* at the origin, the second a *maximum*, while the third has a *saddle point* there.

It is less obvious how to treat polynomials such as

$$w = 3x^2 - 2xy + y^2 \tag{4}$$

or

$$w = 3x^2 - 2xy + y^2 + z^2 \tag{5}$$

with xy terms and possibly more variables. These two examples can be handled if one recognizes that $x^2 - 2xy + y^2 = (x - y)^2$ so, if one makes the change of variables $r = x$, $s = x - y$ and $t = z$ then in the new coordinates these polynomials become

$$w = 2x^2 + (x - y)^2 = 2r^2 + s^2 \quad \text{and} \quad w = 2r^2 + s^2 + t^2,$$

which clearly have minima at the origin in rs and rst space, respectively.

The primary task of this section is to give useful criteria for a quadratic polynomial in several variables to have a maximum, minimum, or saddle point. This will then be used in the next section to generalize the calculus of one variable second derivative test for a local maximum to functions of several variables such as our quadratic polynomials.

The first step is to be a bit more systematic. Rewrite (4) as $w = 3x^2 - xy - yx + y^2$ and observe that using the inner (=dot) product it can be written in the more compact form

$$w = \mathbf{X} \cdot A \mathbf{X}, \tag{6}$$

where

$$\mathbf{X} = \begin{pmatrix} x \\ y \end{pmatrix} \quad \text{and } A \text{ is the } \textit{symmetric} \text{ matrix} \quad A = \begin{pmatrix} 3 & -1 \\ -1 & 1 \end{pmatrix}.$$

Similarly, we can also write

$$w = 3x^2 + 2xy + y^2 - 4yz + xz + 5z^2 \tag{7}$$

$$= 3x^2 + xy + yx + y^2 - 2yz - 2zy + \frac{1}{2}xz + \frac{1}{2}zx + 5z^2 \tag{8}$$

in the form (6) where

$$\mathbf{X} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \quad \text{and } A \text{ is the } \textit{symmetric} \text{ matrix} \quad A = \begin{pmatrix} 3 & 1 & \frac{1}{2} \\ 1 & 1 & -2 \\ \frac{1}{2} & -2 & 5 \end{pmatrix}.$$

In the simplest cases when there are no cross-product terms,

$$w = ax^2 + by^2 + cz^2 \tag{9}$$

the matrix A is a *diagonal* matrix

$$A = \begin{pmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{pmatrix}.$$

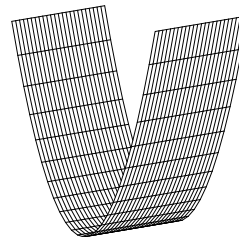
If a, b and c are all positive, then clearly w has a minimum at the origin. Similarly if a, b and c are all negative then clearly w has a maximum at the origin. However if at least one of the coefficients a, b, c is positive and at least one is negative, then the origin is a *saddle point*.

For a given symmetric matrix A we say that the matrix A is

<i>positive definite</i>	if	$\mathbf{X} \cdot \mathbf{A} \mathbf{X} > 0$	for all $\mathbf{X} \neq 0$
<i>negative definite</i>	if	$\mathbf{X} \cdot \mathbf{A} \mathbf{X} < 0$	for all $\mathbf{X} \neq 0$
<i>indefinite</i>	if	$\mathbf{X} \cdot \mathbf{A} \mathbf{X}$ changes sign	

Thus, *positive definite* means precisely that the quadratic polynomial $w = \mathbf{X} \cdot \mathbf{A} \mathbf{X}$ has a minimum only at the origin, *negative definite* means that w has a maximum only at the origin, and *indefinite* means that w has a saddle point at the origin.

If A is invertible, we will see that these three cases are the only possibilities, while if A is not invertible, there are a few degenerate cases not included in the above. One of the simplest examples is the polynomial $w = x^2 + 2xy + y^2$. Since $w = (x + y)^2 \geq 0$, it is almost positive definite. The difficulty is that $w = 0$ at points other than just $x = y = 0$ since it is zero along the whole line $y = -x$. This is clearer from its graph.



For situations such as these, we say that the quadratic polynomial $\mathbf{X} \cdot \mathbf{A} \mathbf{X}$ is *positive semi-definite* if $\mathbf{X} \cdot \mathbf{A} \mathbf{X} \geq 0$ for all \mathbf{X} with the understanding that it may be zero for values of \mathbf{X} other than just zero. There is a similar definition of *negative semi-definite*.

We can now rephrase the primary task for this section: give criteria on the matrix A for the quadratic polynomial $w = \mathbf{X} \cdot \mathbf{A} \mathbf{X}$ to be positive definite, negative definite, or indefinite. We will give two different versions, one using eigenvalues and one using determinants. Both of these have their own virtues and are applicable in different situations.

Before venturing further, we make four useful elementary observations.

OBSERVATION 1. *A diagonal matrix is positive definite if and only if all its diagonal elements are positive, it is negative definite if and only if all its diagonal elements are negative.* We mentioned this above. For a diagonal matrix, the corresponding quadratic polynomial looks like (9), so this assertion should be obvious.

OBSERVATION 2. *If A is positive definite, then $-A$ is negative definite.* The simplest example is $w = x^2 + y^2$ where $A = I$ is the identity matrix. w has a minimum only at the origin. The polynomial $w = -(x^2 + y^2)$, where $A = -I$, has a maximum at the origin.

OBSERVATION 3. *If a symmetric matrix A , has at least one positive diagonal element and at least one negative diagonal element, then it is indefinite.* For instance, the matrix A associated with the quadratic polynomial

$$w = 3x^2 + 2xy + y^2 - 4yz + xz - 5z^2$$

has two positive diagonal elements (the coefficients of x^2 and of y^2), and one negative (the coefficient of z^2). At points of the form $(x, 0, 0)$ the polynomial $w = 3x^2$ is clearly positive, while at points of the form $(0, 0, z)$, $w = -5z^2$ is clearly negative. Thus the origin is neither a max nor a min, it is a saddle point. We conclude that A is *indefinite*.

The same reasoning shows that *the diagonal elements of a positive definite matrix are positive while the diagonal elements of a negative definite matrix are negative.*

OBSERVATION 4. *If the symmetric matrix A is positive definite, then it is invertible.* To see this, note that if A were not invertible, then there would be a vector $\mathbf{Z} \neq 0$ with $\mathbf{A} \mathbf{Z} = 0$. But then $\mathbf{Z} \cdot \mathbf{A} \mathbf{Z} = 0$ which contradicts A being positive definite.

Criterion for Positive Definite Using Eigenvalues

The key ingredients to this approach are OBSERVATION 1 just above and the fundamental fact that one can always diagonalize a symmetric matrix. To be more specific, given a real symmetric $n \times n$ matrix A there is an orthogonal matrix R whose columns are eigenvectors of A so that the matrix $D := R^{-1}AR$ is a diagonal matrix. The diagonal elements of D are the eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$ of A . From this we can write $A = RDR^{-1}$. Thus, using the general rule $\mathbf{X} \cdot R\mathbf{Y} = R^T\mathbf{X} \cdot \mathbf{Y}$ and since $R^T = R^{-1}$, we find the quadratic polynomial is

$$w = \mathbf{X} \cdot A\mathbf{X} = \mathbf{X} \cdot RDR^{-1}\mathbf{X} = R^{-1}\mathbf{X} \cdot DR^{-1}\mathbf{X}.$$

This leads us to make the change of coordinates $\mathbf{V} = R^{-1}\mathbf{X}$. In these $\mathbf{V} = (v_1, \dots, v_n)$ coordinates

$$w = \mathbf{V} \cdot D\mathbf{V} = \lambda_1 v_1^2 + \lambda_2 v_2^2 + \dots + \lambda_n v_n^2. \quad (10)$$

We now apply OBSERVATION 1 and conclude that w has a minimum only at the origin if and only if the eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$ are positive. To summarize:

EIGENVALUE TEST. *A symmetric matrix A is:*

- positive definite *if and only if all of its eigenvalues are positive*
- negative definite *if and only if all of its eigenvalues are negative*
- indefinite *if and only if some eigenvalues are positive and some negative.*

If A is invertible, that is, if none of its eigenvalues are zero, these three are the *only* possibilities. However, some degenerate cases, such as a matrix having some positive eigenvalues and the remaining eigenvalues being zero, are not included in the above. This is treated further in Exercises 26-28.

EXAMPLE 1. The quadratic polynomial

$$4x^2 - 6xy + 4y^2 + 2z^2$$

is associated with the matrix

$$A = \begin{pmatrix} 4 & -3 & 0 \\ -3 & 4 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

whose characteristic polynomial is $(2 - \lambda)[(4 - \lambda)^2 - 9] = (2 - \lambda)(\lambda - 7)(\lambda - 1)$. Thus its eigenvalues are 1, 2 and 7. Since these eigenvalues are all positive, the matrix A is positive definite. Notice that this matrix A does have some negative elements, although by OBSERVATION 3 they couldn't have been on the diagonal since then A could not possibly have been positive definite.

Criterion for Positive Definite Using Determinants

One can also use determinants to test if a matrix is positive definite. This has the advantage that one does not need to find the eigenvalues. Associated with an $n \times n$ matrix A one has the *principal minors*

$$A_1 = (a_{11}), A_2 = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, A_3 = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}, \dots, A_n = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \dots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}.$$

DETERMINANT TEST. *The symmetric matrix A is positive definite if and only if the determinants of all the principal minors are positive. Similarly, by OBSERVATION 2, A is negative definite if and only if these determinants alternate in sign: $\det A_1 < 0$, $\det A_2 > 0$, \dots*

The above assertion concerning negative definite matrices follows from the first part by using OBSERVATION 2.

This criterion is not obvious to most people. However it is easy to use, at least for matrices that are not too large.

EXAMPLE 1 (again). The principal minors are

$$A_1 = (4), \quad A_2 = \begin{pmatrix} 4 & -3 \\ -3 & 4 \end{pmatrix}, \quad \text{and} \quad A_3 = \begin{pmatrix} 4 & -3 & 0 \\ -3 & 4 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

Since $\det A_1 = 4 > 0$, $\det A_2 = 7 > 0$, and $\det A_3 = 14 > 0$, by this determinant criterion the matrix A is positive definite.

EXAMPLE 2. Although the *diagonal* elements of a positive definite matrix must be positive (OBSERVATION 3), the sign of the off-diagonal elements is not significant. For instance, consider the matrices

$$\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}, \quad \text{and} \quad \begin{pmatrix} 1 & -2 \\ -2 & 1 \end{pmatrix}$$

associated with the quadratic polynomials $w = x^2 \pm 4xy + y^2$. By the determinant test both are indefinite so these quadratic polynomials both have saddle points at the origin. One can transform from one of these polynomials to the other merely by replacing y with $-y$, that is, by making a reflection across the x -axis; clearly this does not change the essential behavior of the polynomial.

EXAMPLE 3. The symmetric matrix

$$\begin{pmatrix} 1 & 2 & -2 & 5 & 3 \\ 2 & 9 & 7 & 0 & 8 \\ -2 & 7 & 7 & 0 & 3 \\ 5 & 0 & 0 & 1 & 0 \\ 3 & 8 & 3 & 0 & -5 \end{pmatrix}$$

is indefinite. To see this one could use the determinant test, but it is simpler to use OBSERVATION 3 since at least one of the diagonal elements is positive and a at least one is negative.

PROOF OF THE DETERMINANT TEST. The proof is instructive, but complicated. It is included so the more adventurous readers can see how it works. Upon first reading it is best just to skim it.

PART 1. *If A is positive definite then the determinants of all its principal minors are positive.*

Step 1. We first show that if the symmetric matrix A is positive definite, then its determinant is positive. In fact, we give two different proofs of this.

The first proof begins from the basic fact that the determinant of matrix is the product of its eigenvalues. If A is positive definite, above we showed that all of its eigenvalues must be positive, hence the determinant is positive.

The second proof starts from the observation that the sum of two positive definite symmetric matrices A and B is positive [proof:

$$\mathbf{X} \cdot (A + B)\mathbf{V} = \mathbf{X} \cdot A\mathbf{V} + \mathbf{X} \cdot B\mathbf{V} > 0$$

for all $\mathbf{X} \neq 0$.] In particular, if A is positive definite, so is the matrix $C(t) = tA + (1-t)I$ for $0 \leq t \leq 1$. Since $C(t)$ is positive definite, by OBSERVATION 4 above it is invertible and hence $\det C(t) \neq 0$ for $0 \leq t \leq 1$. Since $\det C(0) = \det I > 0$, then $\det C(1) = \det A > 0$.

Step 2. We claim that if A is positive definite, then so are its principal minors. For this we must show, for instance, that A_3 is positive definite. Let $\mathbf{V} = (v_1, v_2, v_3)$ and let \mathbf{X} be a vector whose first part is \mathbf{V} and the rest is 0, so $\mathbf{X} = (v_1, v_2, v_3, 0, \dots, 0)$. If $\mathbf{V} \neq 0$, then since A is positive definite we know

$$0 < \mathbf{X} \cdot A\mathbf{X} = \mathbf{V} \cdot A_3\mathbf{V}.$$

Thus A_3 is positive definite. One similarly shows that A_k is positive definite for any $1 \leq k \leq n$. Since these A_k are positive definite, by Step 1 their determinants are all positive.

PART II. Conversely, if the determinants of all its principal minors are positive, then A is positive definite.

Step 1. We need a preliminary computation relating $\det A$ and the matrix A_{n-1} . For cleanliness, let $B = A_{n-1}$ and write A as

$$A = \begin{pmatrix} & & & a_{1n} \\ & B & & \vdots \\ & & & a_{n-1n} \\ a_{n1} & \cdots & a_{nn-1} & c \end{pmatrix},$$

where $a_{nn} = c$. If we let α be the column vector $\alpha = (a_{1n}, a_{2n}, \dots, a_{n-1n})$ this can be further abbreviated as

$$A = \begin{pmatrix} B & \alpha \\ \alpha^T & c \end{pmatrix}. \quad (11)$$

We want to compute $\det A$. Perhaps the simplest approach is to notice that matrix multiplication also works for *block matrices* such as (11). In the formula below let I be the $(n-1) \times (n-1)$ identity matrix. We seek a column vector \mathbf{v} with the same shape as α so that the product

$$\begin{pmatrix} B & \alpha \\ \alpha^T & c \end{pmatrix} \begin{pmatrix} I & \mathbf{v} \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} B & B\mathbf{v} + \alpha \\ \alpha^T & \alpha^T\mathbf{v} + c \end{pmatrix} \quad (12)$$

is as simple as possible. Notice that the second matrix on the left is upper triangular, so its determinant is the product of the diagonal elements and hence is 1 no matter how we pick \mathbf{v} . We'll use $B\mathbf{v} + \alpha = 0$; that is, assuming B is invertible, then $\mathbf{v} = -B^{-1}\alpha$. With this choice of \mathbf{v} we can evaluate the determinant on the right easily, expanding by minors using the last column. Thus, we take the determinant of both sides of (12) to conclude that

$$\det A = (c - \alpha^T B^{-1}\alpha) \det B = (c - \alpha \cdot B^{-1}\alpha) \det B. \quad (13)$$

Step 2. We will use this to show that if B is a positive definite symmetric matrix and if A in (11) has $\det A > 0$, then A is also positive definite.

We must show that $\mathbf{X} \cdot A\mathbf{X} > 0$ for all $\mathbf{X} \neq 0$. To utilize the block structure of (11), write \mathbf{X} as the column vector $\mathbf{X} = (\mathbf{x}, z)$ where $\mathbf{x} = (x_1, \dots, x_{n-1})$ and z is a scalar. Then from (11)

$$A\mathbf{X} = \begin{pmatrix} B & \alpha \\ \alpha^T & c \end{pmatrix} \begin{pmatrix} \mathbf{x} \\ z \end{pmatrix} = \begin{pmatrix} B\mathbf{x} + \alpha z \\ \alpha^T\mathbf{x} + cz \end{pmatrix}. \quad (14)$$

Consequently

$$\mathbf{X} \cdot A\mathbf{X} = \mathbf{x} \cdot B\mathbf{x} + 2\alpha \cdot \mathbf{x}z + cz^2. \quad (15)$$

Since B is positive definite, in Part 1 we showed that $\det B > 0$ so B is invertible. Because $\det A > 0$, equation (13) tells us that $c > \alpha \cdot B^{-1} \alpha$. We use this in (15) to find

$$\mathbf{X} \cdot A \mathbf{X} \geq \mathbf{x} \cdot B \mathbf{x} + 2\alpha \cdot \mathbf{x} z + \alpha \cdot B^{-1} \alpha z^2 \quad (> 0 \text{ unless } z = 0), \quad (16)$$

To make this equation clearer, as above let $\mathbf{v} = -B^{-1} \alpha$, so $\alpha = -B \mathbf{v}$. Then we can rewrite (16) as

$$\mathbf{X} \cdot A \mathbf{X} \geq \mathbf{x} \cdot B \mathbf{x} - 2B \mathbf{v} \cdot \mathbf{x} z + B \mathbf{v} \cdot \mathbf{v} z^2 \quad (17)$$

$$= (\mathbf{x} - \mathbf{v} z) \cdot B (\mathbf{x} - \mathbf{v} z). \quad (18)$$

Since the right side of this is of the form $\mathbf{w} \cdot B \mathbf{w}$ and because B is positive definite, we conclude that (17) is positive unless $\mathbf{x} - \mathbf{v} z = 0$ and, from (16), $z = 0$. Thus (17) is positive unless $\mathbf{x} = 0$ and $z = 0$, so it is positive for all $\mathbf{X} \neq 0$. This proves that A is positive definite.

Step 3. We finally show that if the principal minors of A all have positive determinant, then A is positive definite. If $a_{11} > 0$ and $\det A_2 > 0$, then by Step 2, A_2 is positive definite. But since $\det A_3 > 0$ Step 2 again implies that A_3 is positive definite. Repeating this we conclude eventually that $A_n = A$ is positive definite. This completes the proof.

SOME PICTURES.

One way to understand the quadratic polynomial $w = \mathbf{X} \cdot A \mathbf{X}$ is to look at the level sets; thus so we are looking at the points \mathbf{X} where $\mathbf{X} \cdot A \mathbf{X} = k$, for the particular values $k = 1, 0$, and -1 . As we saw in equation (10), by using a rotation, that is, an orthogonal transformation, we need only consider the special case when A is diagonal. For pictures, we will look at the special cases of two and three dimensions.

Two dimensions

$$\lambda_1 x_1^2 + \lambda_2 x_2^2 = c. \quad (19)$$

There are two cases, depending if λ_1 and λ_2 have the same or have opposite sign. If λ_1 and λ_2 have the *same* sign, then since we can always change the sign of c , we make assume that both λ 's are positive and write $\lambda_1 = 1/a^2$, $\lambda_2 = 1/b^2$. Then our equation is

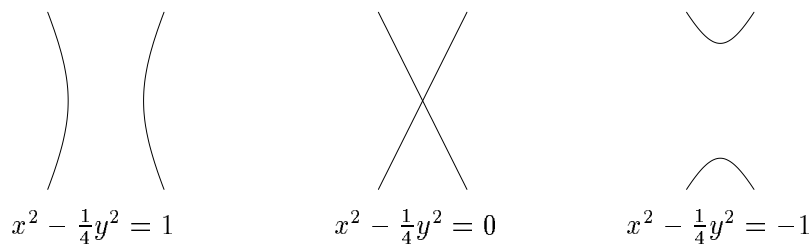
$$\frac{x_1^2}{a^2} + \frac{x_2^2}{b^2} = k.$$

If $k = 1$, this defines an ellipsoid whose semi-axes have lengths a and b , respectively. If $k = 0$, only the origin satisfies the equation, while if $k = -1$ no points satisfy the equation. Of course all of this is obvious from the three dimensional graph of $w = x_1^2/a^2 + x_2^2/b^2$. The alternate approach here will make it easier to compare with the other cases below.

If λ_1 and λ_2 have the *opposite* sign, then say $0 < \lambda_1 = 1/a^2$ and $0 > \lambda_2 = -1/b^2$. The equation becomes

$$\frac{x_1^2}{a^2} - \frac{x_2^2}{b^2} = k.$$

If $k = 1$ this is a hyperbola that opens on the horizontal axis, if $k = 0$ the equation is $(x_1 - x_2)(x_1 + x_2) = 0$ which is the pair of straight lines $x_1 = \pm x_2$. Finally if $k = -1$ this is a hyperbola that opens on the vertical axis.



Three dimensions

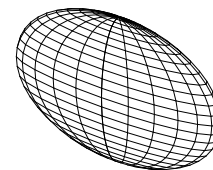
$$\lambda_1 x_1^2 + \lambda_2 x_2^2 + \lambda_3 x_3^2 = 1. \tag{20}$$

The three cases depend on how many of the λ 's are positive or negative.

All λ 's positive. For reasons that will be clear shortly, it is convenient to write $\lambda_1 = 1/a^2$, $\lambda_2 = 1/b^2$, and $\lambda_3 = 1/c^2$, so the equation is

$$\frac{x_1^2}{a^2} + \frac{x_2^2}{b^2} + \frac{x_3^2}{c^2} = k.$$

This is an *ellipsoid*. If $k > 0$ we can always divide by it to reduce to the case where $k = 1$ (the cases $k \leq 0$ are not interesting, Why?). Thus, we assume $k = 1$. Since x_1 is largest when $x_2 = x_3 = 0$, in which case $x_1 = \pm a$, we see that the length of this semi-axis is a . Similarly the lengths of the other semi-axes are b and c , respectively. If $k = 0$ the only point that satisfies this is the origin, while if $k = -1$, no real points satisfy this.



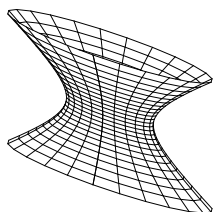
$$x^2 + \frac{1}{25}y^2 + \frac{1}{9}z^2 = 1.$$

Two λ 's positive, one negative. In this case we write, say, $\lambda_1 = a^2$, $\lambda_2 = b^2$, while $\lambda_3 = -c^2$ and the equation is

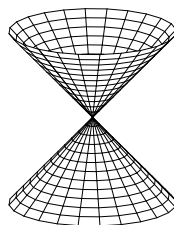
$$\frac{x_1^2}{a^2} + \frac{x_2^2}{b^2} - \frac{x_3^2}{c^2} = k.$$

If $k = 1$ this is a *hyperboloid of one sheet*. If $k = 0$ this is a *cone*, while if $k = -1$ this is a *hyperboloid of two sheets*.

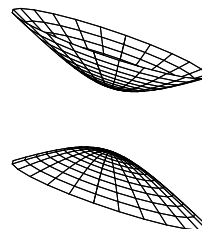
One λ positive, two negative. Multiplying the equation by -1 reduces this case to the previous case (just change the sign of k).



$x^2 + y^2/25 - z^2/9 = 1$
hyperboloid of one sheet



$x^2 + y^2/25 - z^2/9 = 0$
cone



$x^2 + y^2/25 - z^2/9 = -1$
hyperboloid of two sheets

QUADRATIC POLYNOMIALS WITH LOWER ORDER TERMS

We close this section by showing how to treat quadratic polynomials such as

$$w = 3x^2 - 2xy + y^2 + z^2 - 2x + y - 5z + 11$$

with lower order terms. Write this in the form

$$w = \mathbf{X} \cdot A\mathbf{X} + \mathbf{b} \cdot \mathbf{X} + c, \quad (21)$$

where the matrix A is as before, $\mathbf{b} = (-2, 1, -5)$ and $c = 11$. We will use the approach mentioned at the beginning of this section for a polynomial in one variable and make a change of variable $\mathbf{X} = \mathbf{V} + \mathbf{r}$ where \mathbf{r} is to be found. Substituting and computing gives

$$w = (\mathbf{V} + \mathbf{r}) \cdot A(\mathbf{V} + \mathbf{r}) + \mathbf{b} \cdot (\mathbf{V} + \mathbf{r}) + c \quad (22)$$

$$= \mathbf{V} \cdot A\mathbf{V} + (\mathbf{b} + 2A\mathbf{r}) \cdot \mathbf{V} + \mathbf{r} \cdot A\mathbf{r} + \mathbf{b} \cdot \mathbf{r} + c. \quad (23)$$

Now we choose \mathbf{r} to eliminate the linear term in \mathbf{v} ; thus $2A\mathbf{r} = -\mathbf{b}$, so $\mathbf{r} = -\frac{1}{2}A^{-1}\mathbf{b}$. Then w has the desired simpler form

$$w = \mathbf{V} \cdot A\mathbf{V} + k = (\mathbf{X} - \mathbf{r}) \cdot A(\mathbf{X} - \mathbf{r}) + k,$$

where k is the constant

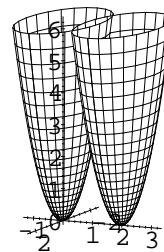
$$k = \mathbf{r} \cdot A\mathbf{r} + \mathbf{b} \cdot \mathbf{r} + c = c - \frac{1}{4}\mathbf{b} \cdot A^{-1}\mathbf{b}.$$

This procedure is the generalization of “completing the square” to quadratic polynomials in several variables.

Just as for a quadratic polynomial in one variable, another way to find \mathbf{r} is to observe from (21) that since $\text{grad } w = 2A\mathbf{X} + \mathbf{b}$, the only critical point of (21) is at $\mathbf{X} = -(1/2)A^{-1}\mathbf{b}$. Thus translating by $(1/2)A^{-1}\mathbf{b}$ places this critical point on the vertical axis of symmetry. Observe that these formulas are essentially identical with (3) for a polynomial in one variable.

We leave it as an exercise for you to find the particular value of r for the special example we have. Note that this procedure presumes that A is invertible, or at least that one can solve $2A\mathbf{r} + \mathbf{b} = 0$.

EXAMPLE. To the right are plots of $z = x^2 + 4y^2$ and its translate, $z = x^2 + 4(y-2)^2$ so the axis of symmetry is the vertical line through $(0, 2, 0)$. Note that for $z = c$, the level curve on the left is the ellipse $x^2 + 4y^2 = c$, while on the right is the same ellipse, just translated: $x^2 + 4(y-2)^2 = c$.



Problems

1. Write the following quadratic polynomials in the form $a(x-r)^2 + k$, that is, find the axis of symmetry r , and the height k .

(a) $x^2 + 4x - 3$

(b) $-2x^2 - 6x + 5$

2. Write each of the following quadratic polynomials in the form $\mathbf{X} \cdot A \mathbf{X}$, that is, find the symmetric matrix A .

- | | |
|------------------------------|---|
| (a) $3x^2 + 3y^2$ | (g) $-2x^2 + xy + y^2 + 4xz - 2yz + 6z^2$ |
| (b) $3x^2 - 4xy - 2y^2$ | (h) $2xy - 6yz + z^2$ |
| (c) $-2x^2 + xy + y^2$ | (i) $(x + y - z + 2u)^2$ |
| (d) $4xy$ | (j) $4xy + y^2 - 7xz + 2z^2$ |
| (e) $x^2 + (x + 2y)^2$ | (k) $x^2 + 2y^2 + 3z^2 - 4w^2 + 4wz$ |
| (f) $x^2 + y^2 + 2xz + 5z^2$ | (l) $x^2 + 2y^2 + 3z^2 + 4w^2 + 6yw$ |

3. Determine which of the quadratic forms in Problem 2 are positive definite (or semi-definite), negative definite (or semi-definite), or indefinite.

4. For the polynomial $3x_1^2 - 4x_1x_2 + 3x_2^2$:

- Determine if it is positive-definite, negative-definite, or indefinite.
- Use Maple to graph it.
- Find an orthogonal matrix R that diagonalizes the symmetric matrix A associated with this quadratic polynomial, so $R^{-1}AR = D$ is a diagonal matrix. Then make the change of variable $\mathbf{X} = R\mathbf{V}$, where $\mathbf{V} = (v_1, v_2)$, and write the polynomial in these new coordinates.

5. Repeat Problem 4. for the polynomial $3x_1^2 + 8x_1x_2 - 3x_2^2$.

6. Repeat Problem 4. for the polynomial $3x_1^2 + 8x_1x_2 - 3x_2^2 - z^2$.

7. Use Maple to graph $3x^2 - 4xy + 3y^2 = 1$. Use all of the three following approaches:

- `plot3d`: $z = x_1^2 - 4x_1x_2 + 3x_2^2$ for $0 \leq z \leq 1$ (Maple uses `view=0..1`), or else show the contour line $z = 1$.
- Plot this curve implicitly.
- Use the procedure of Problem 4c to write this curve as $a^2v_1^2 + b^2v_2^2 = 1$, and then use the parameterization $v_1 = (\cos t)/a$, $v_2 = (\sin t)/b$ for $0 \leq t \leq 2\pi$.

8. Repeat Problem 7 for the quadratic polynomial $3x_1^2 + 8x_1x_2 - 3x_2^2$, only for part c) use hyperbolic functions $\cosh t = \frac{1}{2}(e^t + e^{-t})$ and $\sinh t = \frac{1}{2}(e^t - e^{-t})$. The only property one needs of these hyperbolic functions is the easily verified identity $\cosh^2 t - \sinh^2 t = 1$, which serves as a replacement for $\sin^2 t + \cos^2 t = 1$.

9. Let $q(x, y, r, s) = 5x^2 + 2xy - 6xs + 5y^2 - 6ys - 9s^2 - 18xr - 18yr + 3r^2 + 36rs$. Determine if q is positive definite, negative definite, or indefinite.

10. Let $h(x, y, r, s) = 5x^2 + 2xy - 6xs + 5y^2 - 6ys + 9s^2 - 18xr - 18yr + 3r^2 + 36rs$.

- Determine if h is positive definite, negative definite, or indefinite.
- Find the eigenvalues of the symmetric matrix associated with h . [Remark: Maple does this quite quickly.]

11. Let $p(x, y) = 3x^2 - 4xy + 3y^2 - 14x + 16y + 25$.

- Write this in the form $p(x, y) = (\mathbf{X} - \mathbf{X}_0) \cdot A(\mathbf{X} - \mathbf{X}_0) + c$, that is, find the axis of symmetry \mathbf{X}_0 and the height c . Here, as usual, \mathbf{X} is the *column* vector $\mathbf{X} = (x, y)$.
- Graph this polynomial.

12. Repeat Problem 11 for $q(x, y) = 3x^2 + 8xy - 3y^2 + 24x - 18y + 1$.

13. Let A be a positive definite symmetric matrix.

- Show that A^2 and A^{-1} are also positive definite.
- Better yet, show that A^k is positive definite for *any* integer k .

14. If A is a symmetric matrix, let $f(\mathbf{X}) = \mathbf{X} \cdot A\mathbf{X}$. Use Lagrange multipliers to show that the maximum value of f for all points on the unit ball, $\|\mathbf{X}\| = 1$, is the largest eigenvalue of A .

15. If A is any symmetric matrix, show that there is some constant c so that the matrix $A + cI$ is positive definite. Can you find the optimal value of c ?

16. Let A be a symmetric $n \times n$ matrix with eigenvalues $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$. Show that for any vector \mathbf{X}

$$\lambda_1 \|\mathbf{X}\|^2 \leq \mathbf{X} \cdot A\mathbf{X} \leq \lambda_n \|\mathbf{X}\|^2.$$

[Suggestion: Use equation (10).]

17. Let B be any invertible matrix. Show that $B^T B$ and BB^T are both positive definite. If B is not necessarily invertible – or even a square matrix – show that these are at least positive semi-definite.

18. If A is a symmetric positive definite matrix and C is any invertible matrix, show that $C^T A C$ is also positive definite.

19. A matrix C is called *skew-symmetric* if $C^T = -C$.

- Give an example of a skew-symmetric 2×2 matrix (other than the 0 matrix).
- Show that C is skew-symmetric if and only if $C\mathbf{X} \cdot \mathbf{Y} = -\mathbf{X} \cdot C\mathbf{Y}$ for all vectors \mathbf{X} and \mathbf{Y} .
- If C is skew-symmetric, show that $\mathbf{X} \cdot C\mathbf{X} = 0$ for all vectors \mathbf{X} .

20. Let M be any square matrix and write $M = S + T$, where $S = \frac{1}{2}(M + M^T)$ and $T = \frac{1}{2}(M - M^T)$.

- Show that S is symmetric and T is skew-symmetric.
- Show that if $M = A + B$, where A is symmetric and B is skew symmetric, then in fact $A = S$ and $B = T$.
- Prove the identity $\mathbf{X} \cdot M\mathbf{X} = \mathbf{X} \cdot S\mathbf{X}$ for all vectors \mathbf{X} . Thus, the quadratic polynomial depends only on the symmetric part of M . (This is why we only use symmetric matrices in quadratic polynomials).

21. Let A , B , and C be $n \times n$ symmetric matrices.

- Expand $(\mathbf{X} + \mathbf{Y}) \cdot C(\mathbf{X} + \mathbf{Y})$ to obtain the formula

$$\mathbf{X} \cdot C\mathbf{Y} = \frac{1}{2}[(\mathbf{X} + \mathbf{Y}) \cdot C(\mathbf{X} + \mathbf{Y}) - \mathbf{X} \cdot C\mathbf{X} - \mathbf{Y} \cdot C\mathbf{Y}].$$

The point is that the expressions on the right are all of the form $\mathbf{V} \cdot C\mathbf{V}$ for various vectors \mathbf{V} , while the left side has a more general form.

- Use this to show that if $\mathbf{X} \cdot C\mathbf{X} = 0$ for all vectors \mathbf{X} , then the only possibility is that $C = 0$. [There is an alternate proof by diagonalizing C].
- Show that if the two quadratic polynomials $\mathbf{X} \cdot A\mathbf{X}$ and $\mathbf{X} \cdot B\mathbf{X}$ are equal for all values of \mathbf{X} , then $A = B$. Thus a quadratic polynomial uniquely determines its associated symmetric matrix.

22. Let A be a positive definite matrix.

- Show A has a *square root*, that is, there is a positive definite symmetric matrix P so that $A = P^2$. [Suggestion: First do the special case when A is a diagonal matrix. For the general case, begin by diagonalizing A].

(b) As an example, find the square root of $\begin{pmatrix} 4 & -1 \\ -1 & 4 \end{pmatrix}$

- Let $g(\mathbf{X}) = \mathbf{X} \cdot A\mathbf{X}$. Show that if one makes the change of variable $\mathbf{Y} = P\mathbf{X}$, then g takes the much simpler form $g(\mathbf{X}) = \|\mathbf{Y}\|^2 = y_1^2 + y_2^2 + \dots + y_n^2$. This change of variable is often valuable. Problems 24d, 35, 36 and 37 all benefit from this observation, or variants of it.

- (d) If P is a positive definite square root of A and B is any other symmetric matrix, show that $A - B = P(I - C)P$, where $C = P^{-1}BP^{-1}$ is a symmetric matrix. This is usually more convenient than writing $A - B = A(I - A^{-1}B)$ since C is a symmetric matrix while $A^{-1}B$ is probably not.
23. If M is any invertible matrix, the point of this problem is to show that there is a positive definite matrix P and an orthogonal matrix R so that $M = PR$. This is analogous to writing a complex number z in the polar form $z = re^{i\theta}$. In the complex case, one finds r by the observation that $z\bar{z} = r^2$ and then finds $e^{i\theta}$ using $|z/r| = 1$. The same procedure is used here.
- Observe that if $M = RP$, then $MM^T = P^2$. Since MM^T is positive definite (Problem 17), it has a square root P (Problem 22). This determines P .
 - Now that one knows P , define the matrix R by $R = MP^{-1}$. Show that this R is an orthogonal matrix.
 - Use an analogous procedure to find a positive definite matrix Q and an orthogonal matrix S so that $M = SQ$.
24. For parts a)-b), consider the ellipse $x^2 + y^2/4 = 1$ and the lines $2x + y = c$.
- Use Maple to plot the graphs of the ellipse and the line (on the same plot) for various values of c , both positive and negative.
 - For which value(s) of c does this line intersect the ellipse in exactly one point?
 - Repeat parts a)-b). for the ellipsoid $x^2 + y^2/4 + z^2 = 1$ and, on the same plot, the planes $2x + y + z = c$.
 - More generally, let A be a positive definite symmetric matrix and \mathbf{b} a given non-zero vector. For which value(s) of the constant c does the "plane" $\mathbf{b} \cdot \mathbf{X} = c$ intersect the ellipsoid $\mathbf{X} \cdot \mathbf{A} \mathbf{X} = 1$ in exactly one point? [Suggestion: First do the case when $A = I$, then do the case when A is a diagonal matrix. The answer is $c = \pm\sqrt{\mathbf{b} \cdot A^{-1} \mathbf{b}}$.]
25. Let $\mathbf{V} = (v_1, \dots, v_n)$ be a non-zero column vector and let C be the matrix $C = \mathbf{V}\mathbf{V}^T = (v_i v_j)$, so the j^{th} column of C is $v_j \mathbf{V}$.
- Show that C is positive semi-definite.
 - Compute the inverse of $I + C$.
 - Find the eigenvalues and eigenvectors of C .
 - If A is a symmetric matrix with \mathbf{V} as an eigenvector, find the eigenvalues and eigenvectors of $A + C$. When is this matrix positive definite?
26. Show that both of the following are positive semi-definite, but is not positive definite. Also find the eigenvalues of the associated symmetric matrix, and, for part a), graph $w = q(x, y)$.
- $q(x, y) = x^2 - 6xy + 9y^2$.
 - $Q(x, y, z) = x^2 - 6xy + 9y^2 + z^2$
27. Show that a symmetric matrix is positive semi-definite if and only if all of its eigenvalues satisfy $\lambda_k \geq 0$. State and prove the related assertion for a negative semi-definite matrix.
28. Let A be a symmetric matrix.
- If A is positive semi-definite, show that its principle minors all have non-negative determinant: $\det A_1 \geq 0$, $\det A_2 \geq 0$, etc.
 - Use the example $A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix}$ to show that the converse of part a) is *false*. Do this by showing that $\det A_1 > 0$, $\det A_2 = \det A_3 = 0$ but that A is not positive semi-definite.

29. Let $A = (a_{ij})$ be a symmetric matrix.

- (a) If $a_{11} = 0$ but $a_{12} \neq 0$, show that A is indefinite.
- (b) More generally, if some diagonal element of A is zero but some other element in the *same* column is not zero, show that A is indefinite. This can be thought of as an extension of OBSERVATION 3.

30. Some people define an *orthogonal matrix* R by the property $R^T = R^{-1}$ (transpose is the inverse). The following presents an different approach that is more fundamental.

- (a) If R is an orthogonal matrix, show that for *any* vectors \mathbf{X} and \mathbf{Y} one has

$$R\mathbf{X} \cdot R\mathbf{Y} = \mathbf{X} \cdot \mathbf{Y}.$$

Thus, *orthogonal matrices preserve inner products* – and hence the angles between vectors. As a special case, show that for all vectors \mathbf{X} one has $\|R\mathbf{X}\| = \|\mathbf{X}\|$, that is, *orthogonal matrices preserve the distance from the origin*.

- (b) Conversely, if for *any* vectors \mathbf{X} and \mathbf{Y} one has

$$R\mathbf{X} \cdot R\mathbf{Y} = \mathbf{X} \cdot \mathbf{Y}, \tag{24}$$

show that $R^T = R^{-1}$ (transpose is the inverse). Thus, preserving inner products is equivalent to the other definition of an orthogonal matrix.

- (c) Use the property (24) directly to show that if R and S are both $n \times n$ orthogonal matrices, then so their product, RS .

31. If a matrix R has the property that for *any* vector \mathbf{X} one has $\|R\mathbf{X}\| = \|\mathbf{X}\|$, show that $R^T = R^{-1}$ (transpose is the inverse). Thus, preserving distance to the origin is also equivalent to other definitions of an orthogonal matrix. [Suggestion: use Problem 21a].

32. Let $F(\mathbf{X})$ be a (not necessarily linear) map from R^3 (or R^n) to itself.

- (a) If F has the property that it preserves inner products (see Problem 30. above):

$$F(\mathbf{X}) \cdot F(\mathbf{Y}) = \mathbf{X} \cdot \mathbf{Y}. \tag{25}$$

show that in fact F is linear:

$$F(c\mathbf{X}) = cF(\mathbf{X}) \text{ for any scalar } c, \quad \text{and} \quad F(\mathbf{X} + \mathbf{Y}) = F(\mathbf{X}) + F(\mathbf{Y}) \text{ for any } \mathbf{X} \text{ and any } \mathbf{Y}.$$

Thus F is an orthogonal mapping, $F(\mathbf{X}) = R\mathbf{X}$: [Suggestion: Show that $\|F(\mathbf{X} + \mathbf{Y}) - [F(\mathbf{X}) + F(\mathbf{Y})]\|^2 = 0$, and $\|F(c\mathbf{X}) - cF(\mathbf{X})\|^2 = 0$].

- (b) Show that if a F both fixes the origin and preserves distance between points:

$$F(0) = 0 \quad \text{and} \quad \|F(\mathbf{X}) - F(\mathbf{Y})\| = \|\mathbf{X} - \mathbf{Y}\|,$$

then F has the property (25). Hence it is an orthogonal mapping.

- (c) Let $T(\mathbf{X})$ be a *rigid mapping*, that is, it preserves the distance between any two points:

$$\|T(\mathbf{X}) - T(\mathbf{Y})\| = \|\mathbf{X} - \mathbf{Y}\|.$$

Show that T is the sum of a translation by some vector \mathbf{X}_0 plus an orthogonal transformation, R :

$$T(\mathbf{X}) = \mathbf{X}_0 + R\mathbf{X}.$$

[Suggestion: Apply the preceding part to $F(\mathbf{X}) := T(\mathbf{X}) - T(0)$.]

33. Use both the eigenvalue and determinant tests to decide if the following matrix is positive definite.

$$\begin{pmatrix} 3 & -2 & 3 & -1 & 0 \\ -2 & 6 & -6 & 2 & 0 \\ 3 & -6 & 11 & -3 & 0 \\ -1 & 2 & -3 & 3 & 0 \\ 0 & 0 & 0 & 0 & 2 \end{pmatrix}$$

34. One can think of the vector function $\mathbf{X}(t)$, depending on a parameter t as describing the position of a particle at time t . If

$$\frac{d\mathbf{X}(t)}{dt} = S(t)\mathbf{X},$$

where $S(t)$ is a skew-symmetric matrix, show that the distance from $\mathbf{X}(t)$ to the origin is a constant, that is, $\mathbf{X}(t)$ moves on a sphere of fixed distance from the origin. [Suggestion: Problem 19 is useful here].

35. This problem assumes you know how to make a change of variable in a multiple integral by using the Jacobian determinant. Let $\mathbf{X} = (x_1, \dots, x_n)$.

(a) Using the result from the special case $n = 1$, Evaluate

$$\int_{R^n} e^{-\|\mathbf{X}\|^2} dx_1 \dots dx_n.$$

(b) Let A be a symmetric positive definite $n \times n$ matrix and let

$$Q = \int_{R^n} e^{-\mathbf{X} \cdot A \mathbf{X}} dx_1 \dots dx_n.$$

Use the change of variable in Problem 22c to show that $Q = (\pi/\det A)^{n/2}$.

(c) As an example, apply this to evaluate

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-(x^2 - 2xy + 3y^2)} dx dy.$$

(d) Generalize part a) and evaluate

$$\int_{R^n} e^{-(\mathbf{X} \cdot A \mathbf{X} + 2\mathbf{b} \cdot \mathbf{X} + c)} dx_1 \dots dx_n,$$

where \mathbf{b} is a specified vector and c a scalar.

36. If $A = a_{ij}$ is a positive definite symmetric matrix and $B = (b_{ij})$ is positive semi-definite symmetric (but not identically zero), show that

$$\text{trace}(AB) = \sum_{i,j=1}^n a_{ij}b_{ij} > 0.$$

Here $\text{trace}(A)$ is the sum of the diagonal elements of the square matrix A . [Suggestion: First try the case where $A = I$ and B is a diagonal matrix. For the general case it will help to use the fact that if A and C are similar matrices, then $\text{trace}(A) = \text{trace}(C)$, and also that for any matrices C, D we have $\text{trace}(CD) = \text{trace}(DC)$.]

37. The following problem shows some of the ideas in this section at work in a related setting.

- (a) Let $u(x_1, \dots, x_n)$ be a given function and say one makes the change of variable $\mathbf{Y} = S\mathbf{X}$, where $S = (s_{ij})$ is an invertible matrix. Show that

$$\frac{\partial^2 u}{\partial x_i \partial x_j} = \sum_{k, \ell=1}^n s_{ki} s_{\ell j} \frac{\partial^2 u}{\partial y_k \partial y_\ell}$$

- (b) If $A = (a_{ij})$ is a positive definite 2×2 matrix and $u(x_1, x_2)$ is a given function, let L be the partial differential operator

$$Lu := a_{11} \frac{\partial^2 u}{\partial x_1^2} + 2a_{12} \frac{\partial^2 u}{\partial x_1 \partial x_2} + a_{22} \frac{\partial^2 u}{\partial x_2^2}.$$

Show that one can introduce new coordinates (y_1, y_2) so that in these new coordinates L is the standard *Laplace operator*

$$Lu = \frac{\partial^2 u}{\partial y_1^2} + \frac{\partial^2 u}{\partial y_2^2}.$$

[Suggestion: If $A = (a)$ is a 1×1 matrix and L is the corresponding ordinary differential operator $Lu = au_{xx}$, then under the change of variable $x = \sqrt{a}y$ we get $Lu = u_{yy}$. This leads one to try the change of variable $\mathbf{X} = \sqrt{A}\mathbf{Y}$, so Problem 22 may help here. From another view it may be useful to think of $Lu := \text{trace}(Au'')$]